

## AMENDMENTS TO THE CLAIMS

1. (Currently amended) A computer-implemented method comprising:
  - modifying a query string of characters using a predetermined set of heuristics;
  - performing a character-by-character comparison of the modified query string with at least one known string of characters in a corpus in order to locate an exact match for the modified query string; and
  - responsive to not finding an exact match, performing the following steps in order to locate an equivalent for the modified query string:
    - forming a plurality of sub-string of characters from the query string, the sub-strings having varying lengths such that at least two of the formed sub-strings differ in length; and
    - using an information retrieval technique on the sub-strings formed from the query string to identify a known string of characters equivalent to the query string,
      - wherein the information retrieval technique further comprises:
      - weighting the sub-strings;
      - scoring known strings of characters; and
      - retrieving information associated with the known string having the highest score.
2. (Cancelled)
3. (Currently amended) The method of claim 1 2, further comprising, responsive to the highest score being greater than a first threshold, automatically accepting the known string having the highest score as an exact match.
4. (Currently amended) The method of claim 1 2, further comprising, responsive to the highest score being less than a second threshold and greater than a first threshold, presenting the known string having the highest score to a user for manual confirmation.

5. (Currently amended) The method of claim 1 2, further comprising, responsive to the highest score being less than a second threshold and greater than a third threshold, presenting the known string having the highest score to a user to select the equivalent string.
6. (Previously presented) The method of claim 1, wherein forming a plurality of sub-strings of character comprises successively extending sub-strings based on frequency of occurrence in the modified query string.
7. (Previously presented) The method of claim 1, wherein the query string is selected from the group consisting of a song title, a song artist, an album name, a book title, an author's name, a book publisher, a genetic sequence, and a computer program.
8. (Previously presented) The method of claim 1, wherein the predetermined set of heuristics comprises removing whitespace from the query string.
9. (Previously presented) The method of claim 1, wherein the predetermined set of heuristics comprises removing a portion of the query string.
10. (Previously presented) The method of claim 1, wherein the predetermined set of heuristics comprises replacing a symbol in the query string with an alternate representation for the symbol.
11. (Previously presented) The method of claim 1 further comprising storing a database entry indicating that the query string is an equivalent of the identified known string.
- 12-37. (Cancelled)
38. (Previously presented) The computer-implemented method of claim 1, wherein the length of a sub-string is determined based on one or more character sequences identified

in the modified query string and a corresponding frequency of occurrence for each identified character sequences.

39. (Currently amended) The computer-implemented method of claim 1 2, wherein a weight for a given sub-string is based at least in part on a number of times the sub-string occurs in the query.

40-43. (Cancelled)